Multiple Worlds, One Universal Wave Function

Evan Hubinger

Harvey Mudd College

(2017 Harvey Mudd College Rojansky Writing Award
winner for most outstanding paper in quantum mechanics.)

(Dated: April 30, 2017)

We seek to present and defend the view that the interpretation of quantum mechanics is no more complicated than the interpretation of plate tectonics: that which is being studied is real, and that which the theory predicts is true. The view which holds that the mathematical formalism of quantum mechanics—without any additional postulates—is a complete description of reality is known as the Everett interpretation. We seek to defend the Everett interpretation of quantum mechanics as the most probable interpretation available. To accomplish this task, we analyze the history of the Everett interpretation, provide mathematical backing for its assertions, respond to criticisms that have been leveled against it, and compare it to its modern alternatives.

I. INTRODUCTION

One of the most puzzling aspects of quantum mechanics is the fact that, when one measures a system in a superposition of multiple states, it is only ever found in one of them. This puzzle was dubbed the "measurement problem," and the first attempt at a solution was by Werner Heisenberg, who in 1927 proposed his theory of "wave function collapse."[1] Heisenberg proposed that there was a cutoff length, below which systems were governed by quantum mechanics, and above which they were governed by classical mechanics. Whenever quantum systems encounter the cutoff point, the theory stated, they collapse down into a single state with probabilities following the squared amplitude, or Born, rule. Thus, the theory predicted that physics just behaved differently at different length scales. This traditional interpretation of quantum mechanics is usually referred to as the Copenhagen interpretation.

From the very beginning, the Copenhagen interpretation was seriously suspect. Albert Einstein was famously displeased with its lack of determinism, saying "God does not play dice," to which Niels Bohr quipped

in response, "Einstein, stop telling God what to do."[2] As clever as Bohr's answer is, Einstein—with his famous physical intuition—was right to be concerned. Though Einstein favored a hidden variable interpretation[3], which was later ruled out by Bell's theorem[4], the Copenhagen interpretation nevertheless leaves open many questions. If physics behaves differently at different length scales, what is the cutoff point? What qualifies as a wave-function-collapsing measurement? How can physics behave differently at different length scales, when macroscopic objects are made up of microscopic objects? Why is the observer not governed by the same laws of physics as the system being observed? Where do the squared amplitude Born probabilities come from? If the physical world is fundamentally random, how is the world we see selected from all the possibilities? How could one explain the applicability of quantum mechanics to macroscopic systems, such as Chandrasekhar's insight in 1930 that modeling neutron stars required the entire star to be treated as a quantum system?[5]

II. THE EVERETT INTERPRETATION OF QUANTUM MECHANICS

Enter the Everett Interpretation. In 1956, Hugh Everett III, then a doctoral candidate at Princeton, had an idea: if you could find a way to explain the phenomenon of measurement from within wave mechanics, you could do away with the extra postulate of wave function collapse, and thus many of the problems of the Copenhagen interpretation. Everett worked on this idea under his thesis advisor, Einstein-prize-winning theoretical physicist John Wheeler, who would later publish a paper in support of Everett's theory.[6] In 1957, Everett finished his thesis "The Theory of the Universal Wave Function,"[7] published as the "'Relative State' Formulation of Quantum Mechanics."[8] In his thesis, Everett succeeded in deriving every one of the strange quirks of the Copenhagen interpretation—wave function collapse, the apparent randomness of measurement, and even the Born rule—from purely wave mechanical grounds, as we will do in section III.

Everett's derivation relied on what was at the time a controversial application of quantum mechanics: the existence of wave functions containing observers themselves. Everett believed that there was no reason to re-

strict the domain of quantum mechanics to only small, unobserved systems. Instead, Everett proposed that any system, even the system of the entire universe, could be encompassed in a single, albeit often intractable, "universal wave function."

Modern formulations of the Everett interpretation reduce his reasoning down to two fundamental ideas:[9][10][11][12][13]

1. the wave function obeys the standard, linear, deterministic Schrodinger wave equation *at all times* (the relativistic variant, to be precise), and

2. the wave function is *physically real.*

Specifically, the first statement precludes wave function collapse and demands that we continue to use the same wave mechanics for all systems, even those with observers, and the second statement demands that we accept the physical implications of doing so. The Everett interpretation is precisely that which is implied by these two statements.

Importantly, neither of these two principles are additional assumptions on top of traditional quantum theory—instead, they are simplifications of existing quantum theory, since they act only to remove the prior ad-hoc postulates of wave function collapse and the non-universal applicability of the wave equation.[11][14] The beauty of the Everett interpretation is the fact that we can remove the postulates of the Copenhagen interpretation and still end up with a theory that works.

A. DeWitt's Multiple Worlds

Removing the Copenhagen postulates had some implications that did not mesh well with many physicists' existing physical intuitions. If one accepted Everett's universal wave function, one was forced to accept the idea that macroscopic objects—cats, people, planets, stars, galaxies, even the entire universe—could be in a superposition of many states, just as microscopic objects could. In other words, multiple different versions of the universe—multiple worlds, so to speak—could exist simultaneously. It was for this reason that Einstein-prize-winning physicist Bryce DeWitt, a supporter of the Everett interpretation, dubbed Everett's theory of the universal wave function the "multiworld" (or now more commonly "multiple worlds") interpretation of quantum mechanics.[9]

While the idea of multiple worlds may at first seem strange, to Everett, it was simply an extension of the normal laws of quantum mechanics. Simultaneous superposition of states is something physicists already accept for microscopic systems whenever they do quantum mechanics—by virtue

of the overwhelming empirical evidence in favor of it. Not only that, but evidence keeps coming out demonstrating superpositions at larger and larger length scales. In 1999 it was demonstrated, for example, that Carbon-60 molecules can be put into a superposition.[15]. While it is unlikely that a superposition of such a macroscopic object as Schrodinger's cat will ever be conclusively demonstrated, due to the difficulty in isolating such a system from the outside world, it is likely that the trend of demonstrating superposition at larger and larger length scales will continue. It seems that to not accept that a cat could be in a superposition, even if we can never demonstrate it, however, is a failure of induction—a rejection of an empirically-demonstrated trend.

While the Everett interpretation ended up implying the existence of multiple worlds, this was never Everett's starting point. The "multiple worlds" of the Everett interpretation were not added to traditional quantum mechanics as new postulates, but rather fell out from the act of *taking away* the existing ad-hoc postulates of the Copenhagen interpretation—a consequence of taking the wave function seriously as a fundamental physical entity. In Everett's own words, "The aim is not to deny or contradict the conventional formulation of quantum theory, which has demonstrated its usefulness in an overwhelming variety of problems, but rather to supply a new, more general and complete formulation, from which the conventional interpretation can be *deduced.*"[8] Thus, it is not surprising that Stephen Hawking and Nobel laureate Murray Gell-Mann, supporters of the Everett interpretation, have expressed reservations with the name "multiple worlds interpretation," and therefore we will continue to refer to the theory simply as the Everett interpretation instead.[16]

B. The Nature of Observation

Accepting the Everett interpretation raises an important question: if the macroscopic world can be in a superposition of multiple states, what differentiates them? Stephen Hawking has the answer: "in order to determine where one is in space-time one has to measure the metric and this act of measurement places one in one of the various different branches of the wave function in the Wheeler-Everett interpretation of quantum mechanics."[17] When we perform an observation on a system whose state is in a superposition of eigenfunctions, a version of us sees each different, possible eigenfunction. The different worlds are defined by the different eigenfunctions that are observed.

We can show this, as Everett did, just by acknowledging the existence of universal,

joint system-observer wave functions.[7][8] Before measuring the state of a system in a superposition, the observer and the system are independent—we can get their joint wave function simply by multiplying together their individual wave functions. After measurement, however, the two become entangled—that is, the state of the observer becomes dependent on the state of the system that was observed. The result is that for each eigenfunction in the system's superposition, the observer's wave function evolves differently. Thus, we can no longer express their joint wave function as the product of their individual wave functions. Instead, we are forced to express the joint wave function as a sum of different components, one for each possible eigenfunction of the system that could be observed. These different components are the different "worlds" of the Everett interpretation, with the only difference between them being which eigenfunction of the system was observed. We will formalize this reasoning in section III A.

We are still left with the question, however, of why we experience a particular probability of seeing some states over others, if every state that can be observed is observed. Informally, we can think of the different worlds—the different possible observations—as being "weighted" by their squared amplitudes, and which one of the versions of us we are as a random choice from that weighted distribution. Formally, we can prove that under the Everett interpretation, if an observer interacts with many systems each in a superposition of multiple states, the distribution of states they see will follow the Born rule.[7][8][18][11][19][14] A portion of Everett's proof of this fact is included in section III C.

III. THE MATHEMATICS OF THE EVERETT INTERPRETATION

Previously, we asserted that universally-applied wave mechanics was sufficient, without ad-hoc postulates such as wave function collapse, to imply all the oddities of the Copenhagen interpretation. We will now prove that assertion. In this section, as per the Everett interpretation, we will accept that basic wave mechanics is obeyed for all physical systems, including those containing observers. From that assumption, we will show that the apparent phenomena of wave function collapse, random measurement, and the Born Rule follow. The proofs given below are adopted from Everett's original paper.[7][8]

A. The Apparent Collapse of The Wave Function

Suppose we have a system S with eigenfunctions $\{\varphi_i\}$ and initial state $\varphi = \sum_i a_i\varphi_i$. Consider an observer O with initial state ψ. Let $\psi_{i,j,\ldots}$ be the state of O after observing eigenfunctions $\varphi_i, \varphi_j, \ldots$ of S. Since we would like to demonstrate how repeated measurements see a collapsed wave function, we will assume that repeated measurement is possible, and thus that the states φ_i of S remain unchanged after observation. As we are working under the Everett interpretation, we will let ourselves define a joint system-observer wave function Ψ with initial configuration

$$\Psi_0 = \psi\varphi = \psi\sum_i a_i\varphi_i$$

Then, our goal is to understand what happens to Ψ when O repeatedly observes S. Thus, we will define Ψ_n to represent the state of Ψ after $n \in \mathbb{N}$ independent observations of S are performed by O.

Consider the simple case where $\varphi = \varphi_0$ and thus we are in initial state $\Psi_0 = \psi\varphi_0$. In this case, by our previous definition of ψ_i and requirement that φ_i remain unchanged, we can write the state after the observation as $\Psi_1 = \psi_0\varphi_0$. Since quantum mechanics is linear, and the eigenfunctions φ_i are orthogonal, it must be that this same process occurs for each φ_i.

Thus, by the principle of superposition, we can write Ψ_1 in its general form as

$$\Psi_1 = \sum_i a_i\psi_i\varphi_i$$

For the next observation, each ψ_i will once again see the same φ_i, since it has not changed state. As previously defined, we use the notation $\psi_{i,i}$ to denote the state of O after observing S in state φ_i twice. Thus, we can write Ψ_2 as

$$\Psi_2 = \sum_i a_i\psi_{i,i}\varphi_i$$

and more generally, we can write Ψ_n as

$$\Psi_n = \sum_i a_i \psi_{i,i,\ldots,i} \varphi_i$$

where i is repeated n times in $i, i, \ldots, i$.

Thus, once a measurement of S has been performed, every subsequent measurement will see the same eigenfunction, even though all eigenfunctions continue to exist. We can see this from the fact that the same i is repeated in each state $\psi_{i,i,\ldots,i}$ of O. In this way, we see how, despite the fact that the original wave function $\varphi = \sum_i a_i \varphi_i$ for S is in a superposition of many eigenfunctions, once a measurement has been performed, each subsequent measurement will always see the same eigenfunction.

Note that there is no longer a single, independent state ψ of O. Instead, there are many $\psi_{i,i,\ldots,i}$, one for each eigenfunction. What does that mean? It means that for every eigenfunction φ_i of S, there is a corresponding state $\psi_{i,i,\ldots,i}$ of O wherein O sees that eigenfunction. Thus, one is required to accept that there are many observers O_i, with corresponding state $\psi_{i,i,\ldots,i}$, each one seeing a different eigenfunction φ_i. This is the origin of the Everett interpretation's "multiple worlds."

From the perspective of each O_i in this scenario it will appear as if φ has "collapsed" from a complex superposition $\sum_i a_i \varphi_i$ into a single eigenfunction φ_i. As we can see from the joint wave function, however, that is not the case—in fact, the entire superposition still exists. What has changed is only that ψ, the state of O, is no longer independent of that superposition, and has instead become entangled with it.

B. The Apparent Randomness of Measurement

Suppose we now have many such systems S, which we will denote S_n where $n \in \mathbb{N}$. Consider O from before, but with the modification that instead of repeatedly observing a single S, O observes different S_n in each measurement, such that Ψ_n is the joint system-observer wave function after measuring the nth S_n.

As before, we will define the initial joint wave function Ψ_0 as

$$\Psi_0 = \psi \sum_{i_1,i_2,\ldots,i_n} \Big(a_{i_1,i_2,\ldots,i_n} \varphi_{i_1}(x_1)\varphi_{i_2}(x_2)\cdots\varphi_{i_n}(x_n) \Big)$$

where we are summing over all possible combinations of eigenfunctions for the different systems S_n with arbitrary coefficients $a_{i_1,i_2,\ldots,i_n}$ for each combination.

Then, as before, we can use the principle of superposition to find Ψ_1 as

$$\Psi_1 = \sum_{i_1,i_2,\ldots,i_n} \Big(\psi_{i_1} a_{i_1,i_2,\ldots,i_n} \varphi_{i_1}(x_1)\varphi_{i_2}(x_2)\cdots\varphi_{i_n}(x_n) \Big)$$

since the first measurement will see the state φ_{i_1} of S_1. More generally, we can write Ψ_n as

$$\Psi_n = \sum_{i_1,i_2,\ldots,i_n} \Big(\psi_{i_1,i_2,\ldots,i_n} a_{i_1,i_2,\ldots,i_n} \varphi_{i_1}(x_1)\varphi_{i_2}(x_2)\cdots\varphi_{i_n}(x_n)\Big)$$

following the same principle, as each measurement of an S_n will see the corresponding state φ_{i_n}.

Thus, when subsequent measurements of identical systems S_n are performed, the resulting sequence of eigenfunctions observed by O in each ψ appear random (according to what distribution we will show in the next subsection), since there is no structure to the sequences $i_1, i_2, \ldots, i_n$. This appearance of randomness is true even though the entire process is completely deterministic. If, alternatively, O was to return to a previously-measured S_n, we would get a repeat of the first analysis, wherein O would always see the same state as was previously measured.

C. The Born Probability Rule

As before, consider a system S in state $\sum_i a_i\varphi_i$. To be able to talk about a probability for an observer O to see state φ_i, we need some function $P(a_i)$ that will serve as a measure of that probability.

Since we know that quantum mechanics is invariant up to an overall phase, we will impose the condition on P that it must satisfy the equation

$$P(a_i) = P\left(\sqrt{a_i^* a_i}\right) = P(|a_i|)$$

Furthermore, by the linearity of quantum mechanics, we will impose the condition on P such that for $a\varphi$ defined as

$$a\varphi = \sum_i a_i\varphi_i$$

P must satisfy the equation

$$P(a) = \sum_i P(a_i)$$

Together, these two conditions fully specify what function P must be. Assuming φ is normalized, such that $\sum_i \varphi_i^*\varphi_i = 1$, it must be that

$$a^* a = \sum_i a_i^* a_i$$

or equivalently

$$|a| = \sqrt{\sum_i |a_i|^2}$$

such that

$$P(|a|) = P\left(\sqrt{\sum_i |a_i|^2}\right)$$

which, using the phase invariance condition that $P(|a|) = P(a)$, gives

$$P(a) = P\left(\sqrt{\sum_i |a_i|^2}\right)$$

Then, from the linearity condition, we have

$$P(a) = \sum_i P(a_i)$$

which, by the phase invariance condition, is equivalent to

$$P(a) = \sum_i P\left(\sqrt{|a_i|^2}\right)$$

Putting it all together, we get

$$P(a) = P\left(\sqrt{\sum_i |a_i|^2}\right) = \sum_i P\left(\sqrt{|a_i|^2}\right)$$

then, defining a new function $g(x) = P(\sqrt{x})$, yields

$$g\left(\sum_i |a_i|^2\right) = \sum_i g\left(|a_i|^2\right)$$

which implies that g must be a linear function such that for some constant c

$$g(x) = cx$$

Therefore, since $P(x) = g(x^2)$,

$$P(x) = cx^2$$

which, imposing the phase invariance condition, becomes

$$P(x) = c\,|x|^2$$

which, where c is normalized to 1, is the Born rule.

The fact that this measure is a probability, beyond that it is the only measure that can be, is deserving of further proof. The concept of probability is notoriously hard to define, however, and without a definition of probability, it is just as meaningful to call P something as arbitrary as the "stallion" of the wave function as the "probability." Nevertheless, for nearly every reasonable probability theory that exists, such proofs have been provided. Everett provided a proof based on the standard frequentist definition of probability[7][8], David Deutsch (Oxford theoretical physicist) has provided a proof based on game theory[18], and David Wallace (USC theoretical physicist) has provided a proof based on decision theory[11]. For any reasonable definition of probability, wave mechanics is able to show that the above measure satisfies it in the limit without any additional postulates.[19][14][20]

IV. ARGUMENTS FOR AND AGAINST THE EVERETT INTERPRETATION

> "Despite the unrivaled empirical success of quantum theory, the very suggestion that it may be *literally true as a description of nature* is still greeted with cynicism, incomprehension, and even anger."[21]
>
> David Deutsch, 1996

A. Falsifiability and Empiricism

Perhaps the most common criticism of the Everett interpretation is the claim that it is not falsifiable, and thus falls outside the realm of empirical science.[22] In fact, this claim is simply not true—many different methods for testing the Everett interpretation have been proposed, and, a great deal of empirical data regarding the Everett interpretation is already available.

One such method we have already discussed: the Everett interpretation removes the Copenhagen interpretation's postulate that the wave function must collapse at a particular length scale. Were it ever to be conclusively demonstrated that superposition was impossible past some point, the Everett interpretation would be disproved. Thus, every demonstration performed of superposition at larger and larger length scales—such as for Carbon 60 as was previously mentioned[15]—is a test of the Everett interpretation. Arguably, it is the Copenhagen interpretation which is unfalsifiable, since it makes no claim about where the boundary lies at which wave function collapse occurs, and thus proponents can respond to the evidence of larger superpositions simply by changing their theory and moving their proposed boundary up.

Another method of falsification regards the interaction between the Everett interpretation and quantum gravity. The Everett interpretation makes a definitive prediction that gravity must be quantized. Were gravity not quantized—not wrapped up in the wave function like all the other forces—and instead simply a background metric for the entire wave function, we would be able to detect the gravitational impact of the other states we were in a superposition with.[10][23] In 1957, Richard Feynman, who would later come to explicitly support the Everett interpretation[16] as well as become a Nobel laureate, presented an early version of the above argument as a reason to believe in quantum gravity, arguing, "There is a bare possibility (which I shouldn't mention!) that quantum mechanics fails and becomes classical again when the amplification gets far

enough [but] if you believe in quantum mechanics up to any level then you have to believe in gravitational quantization."[24]

Another proposal concerns differing probabilities of finding ourselves in the universe we are in depending on whether the Everett interpretation holds or not. If the Everett interpretation is false, and the universe only has a single state, there is only one state for us to find ourselves in, and thus we would expect to find ourselves in an approximately random universe. On the other hand, if the Everett interpretation is true, and there are many different states that the universe is in, we could find ourselves in any of them, and thus we would expect to find ourselves in one which was more disposed than average towards the existence of life. Approximate calculations of the relative probability of the observed universe based on the Hartle-Hawking boundary condition strongly support the Everett interpretation.[10]

Finally, as we made a point of being clear about in section II, the Everett interpretation is simply a consequence of taking the wave function seriously as a physical entity. Thus, it is somewhat unfair to ask the Everett interpretation to achieve falsifiability independently of the theory—quantum mechanics—which implies it.[22] If a new theory were proposed that said quantum mechanics stopped working outside of the future light cone of Earth, we would not accept it as a new physical controversy—we would say that, unless there is incredibly strong proof otherwise, we should by default assume that the same laws of physics apply everywhere. The Everett interpretation is just that default—it is only by historical accident that it happened to be discovered after the Copenhagen interpretation. Thus, to the extent that one has confidence in the universal applicability of the principles of quantum mechanics, one should have equal confidence in the Everett interpretation, since it is a logical consequence. It is in fact all the more impressive—and tantamount to its importance to quantum mechanics—that the Everett interpretation manages to achieve falsifiability and empirical support despite its primary virtue of simply saying that quantum mechanics be applied universally.

B. Simplicity

Another common objection to the Everett interpretation is that it "postulates too many universes," which Sean Carroll, a Caltech cosmologist and supporter of the Everett interpretation, calls "the basic silly objection."[25] At this point, it should be very clear why this objection is silly: the Everett interpretation postulates no such thing—the existence of "many universes" is an *implication,* not a *pos-*

tulate, of the theory. Opponents of the Everett interpretation, however, have accused it of a lack of simplicity on the grounds that adding in all those additional universes is unnecessary added complexity, and since by the principle of Occam's razor the simplest explanation is probably correct, the Everett interpretation can be rejected.[26]

In fact, Occam's razor is an incredibly strong argument in favor of the Everett interpretation. To explain this, we will first need to formalize what we mean by Occam's razor, which will require some measure of theoretical computer science. Specifically, we will make use of Solomonoff's theory of inductive inference: the best, most general framework we have for comparing the probability of empirically indistinguishable physical theories.[27][28][29][1] To use Solomonoff's formalism, only one assumption is required of us: under some encoding scheme, competing theories of the universe can be modeled as programs. This assumption does not imply that the universe must be computable, only that it can be computably described, which all physical theories capable of being written down must abide by. From this assumption, and the axioms of probability theory, Solomonoff induction can be derived.[27]

Solomonoff induction tells us that, if we have a set of programs[2] $\{T_i\}$ which encode for empirically indistinguishable physical theories, the probability $\mathbb{P}$ of the theory described by a given program T_i with length in bits (0s and 1s) $|T_i|$ is given by

$$\mathbb{P}(T_i) \sim 2^{-|T_i|}$$

up to a constant normalization factor calculated across all the $\{T_i\}$ to make the probabilities sum to 1.[27] We can see how this makes intuitive sense, since if we are predicting an arbitrary system, and thus have no information about the correctness of a program implementing a theory other than its length in bits, we are forced to assign equal probability to each of the two options for each bit, 0 and 1, and thus each additional bit adds a factor of $\frac{1}{2}$ to the total probability of the program. Furthermore, we can see how Solomonoff induction serves as a formalization of Occam's razor, since it gives us a way of calculating how much to discount longer, more complex theories in favor of shorter, simpler ones.

Now, we will attempt to apply this formalism to assign probabilities to competing interpretations of quantum mechanics, which we will represent as elements of the set $\{T_i\}$. Let W be the shortest program which computes the wave equation. Since the wave equation is a component of all quantum the-

[1] In some of these sources, the equivalent formalism of Kolmogorov complexity is used instead.

[2] To be precise, these should be universal Turing machine programs.

ories, it must be that $|W| \leq |T_i|$. Thus, the smallest that any T_i could possibly be is $|W|$, such that any T_i of length $|W|$ is at least twice as probable as a T_i of any other length. The Everett interpretation is such a T_i, since it requires nothing else beyond wave mechanics, and follows directly from it. Therefore, from the perspective of Solomonoff induction, the Everett interpretation is provably optimal in terms of program length, and thus also in terms of probability.

To get a sense of the magnitude of these effects, we will attempt to approximate how much less probable the Copenhagen interpretation is than the Everett interpretation. We will represent the Copenhagen interpretation C as made of three parts: W, wave mechanics; O, a machine which determines when to collapse the wave function; and L, classical mechanics. Then, where the Everett interpretation E is just W, we can write their relative probabilities as

$$\frac{\mathbb{P}(C)}{\mathbb{P}(E)} = \frac{2^{-|W|-|O|-|L|}}{2^{-|W|}} = 2^{-|O|-|L|}$$

How large are O and L? As a quick Fermi estimate for L, we will take Newton's three laws of motion, Einstein's general relativistic field equation, and Maxwell's four equations of electromagnetism as the principles of classical mechanics, for a total of 8 fundamental equations. Assume the minimal implementation for each one averages 100 bits—a very modest estimate, considering the smallest Chess program ever written is 3896 bits long.[30] Then, the relative probability is at most

$$\frac{\mathbb{P}(C)}{\mathbb{P}(E)} = 2^{-|O|-|L|} < 2^{-|L|} \approx 2^{-800} \approx 2 \cdot 10^{-241}$$

which is about the probability of picking four random atoms in the universe and getting the same one each time, and is thus so small as to be trivially dismissible.

C. The Arrow of Time

Another objection to the Everett interpretation is that it is time-symmetric. Since the Everett interpretation is just the wave equation, its time symmetry follows from the fact that the Schrodinger equation is time-reversal invariant, or more technically, charge-parity-time-reversal (CPT) invariant. The Copenhagen interpretation, however, is not, since wave function collapse is a fundamentally irreversible event.[31] In fact, CPT symmetry is not the only natural property that wave function collapse lacks that the Schrodinger equation has—wave function collapse breaks linearity, unitarity, differentiability, locality, and determinism.[13][12][16][32] The Everett interpretation, by virtue of consisting of nothing but the Schrodinger equation, preserves all of these properties. This is an argu-

ment in favor of the Everett interpretation, since there are strong theoretical and empirical reasons to believe that such symmetries are properties of the universe.[33][34][35][5]

Nevertheless, as mentioned above, it has been argued that the Copenhagen interpretation's breaking of CPT symmetry is actually a point in its favor, since it supposedly explains the arrow of time, the idea that time does not behave symmetrically in our everyday experience.[31] Unfortunately for the Copenhagen interpretation, wave function collapse does not actually imply any of the desired thermodynamic properties of the arrow of time.[31] Furthermore, under the Everett interpretation, the arrow of time can be explained using the standard thermodynamic explanation that the universe started in a very low-entropy state.[36]

In fact, accepting the Everett interpretation gets rid of the need for the current state of the universe to be dependent on subtle initial variations in that low-entropy state.[36] Instead, the current state of the universe is simply one of the many different components of the wave function that evolved deterministically from that initial state. Thus, the Everett interpretation is even simpler—from a Solomonoff perspective—than was shown in section IV B, since it forgoes the need for its program to specify a complex initial condition for the universe with many subtle variations.

V. OTHER INTERPRETATIONS OF QUANTUM MECHANICS

> "The mathematical formalism of the quantum theory is capable of yielding its own interpretation."[9]
>
> Bryce DeWitt, 1970

A. Decoherence

It is sometimes proposed that wave mechanics alone is sufficient to explain the apparent phenomenon of wave function collapse without the need for the Everett interpretation's multiple worlds. The justification for this assertion is usually based on the idea of decoherence. Decoherence is the mathematical result, following from the wave equation, that tightly-interacting superpositions tend to evolve into non-interacting superpositions.[37][38] Importantly, decoherence does not destroy the superposition—it merely "diagonalizes" it, which is to say, it removes the interference terms.[37] After decoherence, one is always still left with a superposition of multiple states.[39][40] The only way to remove the resulting superposition is to assume wave function collapse, which every statistical theory claiming to do away

with multiple worlds has been shown to implicitly assume.[41][19] There is no escaping the logic presented in section III A—if one accepts the universal applicability of the wave function, one must accept the multiple worlds it implies.

That is not to say that decoherence is not an incredibly valuable, useful concept for the interpretation of quantum mechanics, however. In the Everett interpretation, decoherence serves the very important role of ensuring that macroscopic superpositions—the multiple worlds of the Everett interpretation—are non-interacting, and that each one thus behaves approximately classically.[41][40] Thus, the simplest decoherence-based interpretation of quantum mechanics is in fact the Everett interpretation. From the Stanford Encyclopedia of Philosophy, "Decoherence as such does not provide a solution to the measurement problem, at least not unless it is combined with an appropriate interpretation of the theory [and it has been suggested that] decoherence is most naturally understood in terms of Everett-like interpretations."[39] The discoverer of decoherence himself, German theoretical physicist Heinz-Dieter Zeh, is an ardent proponent of the Everett interpretation.[42][36]

Furthermore, we have given general arguments in favor of the existence of the multiple worlds implied by the Everett interpretation, which are all reasons to favor the Everett interpretation over any single-world theory. Specifically, calculations of the probability of the current state of the universe support the Everett interpretation[10], as does the fact that the Everett interpretation allows for the initial state of the universe to be simpler[36].

B. Consistent Histories

The consistent histories interpretation of quantum mechanics, owing primarily to Robert Griffiths, eschews probabilities over "measurement" in favor of probabilities over "histories," which are defined as arbitrary sequences of events.[43] Consistent histories provides a way of formalizing what classical probabilistic questions make sense in a quantum domain and which do not—that is, which are consistent. Its explanation for why this consistency always appears at large length scales is based on the idea of decoherence, as discussed above.[43][44] In this context, consistent histories is a very useful tool for reasoning about probabilities in the context of quantum mechanics, and for providing yet another proof of the natural origin of the Born rule.

Proponents of consistent histories claim that it does not imply the multiple worlds of the Everett interpretation.[43] However, since the theory is based on decoherence, there

are always multiple different consistent histories, which cannot be removed via any natural history selection criterion.[45][44] Thus, just as the wave equation implies the Everett interpretation, so too does consistent histories. To see this, we will consider the fact that consistent histories works because of Feynmann's observation that the amplitude of any given final state can be calculated as the sum of the amplitudes along all the possible paths to that state.[44][46] Importantly, we know that two different histories—for example, the different branches of a Mach-Zender interferometer—can diverge and then later merge back together and interfere with each other. Thus, it is not in general possible to describe the state of the universe as a *single* history, since other, parallel histories can interfere and change how that state will later evolve. A history is great for describing how a state came to be, but not very useful for describing how it might evolve in the future. For that, including the other parallel histories—the full superposition—is necessary.

Once one accepts that the existence of multiple histories is necessary on a microscopic level, their existence on a macroscopic level follows—excluding them would require an extra postulate, which would make consistent histories equivalent to the Copenhagen interpretation. If such an extra postulate is not made, then the result is macroscopic superposition, which is to say, the Everett interpretation. This formulation of consistent histories without any extra postulates has been called the theory of "the universal path integral," exactly mirroring Everett's theory of the universal wave function.[46] The theory of the universal wave function—the Everett interpretation—is to the theory of the universal path integral as wave mechanics is to the sum-over-paths approach, which is to say that they are both equivalent formalisms with the same implications.

C. Pilot Wave Theory

The pilot wave interpretation, otherwise known as the de Broglie-Bohm interpretation, postulates that the wave function, rather than being physically real, is a background which "guides" otherwise classical particles.[47] As we saw with the Copenhagen interpretation, the obvious question to ask of the pilot wave interpretation is whether its extra postulate—in this case adding in classical particles—is necessary or useful in any way. The answer to this question is a definitive no. Heinz-Dieter Zeh says of the pilot wave interpretation, "Bohm's pilot wave theory is successful only because it keeps Schrodinger's (exact) wave mechanics unchanged, while the rest of it is obser-

vationally meaningless and solely based on classical prejudice."[42] As we have previously shown in section III, wave mechanics is capable of solving all supposed problems of measurement without the need for any additional postulates. While it is true that pilot wave theory solves all these problems as well, it does so not by virtue of its classical add-ons, but simply by virtue of including the entirety of wave mechanics.[42][48]

Furthermore, since pilot wave theory has no collapse postulate, it does not even get rid of the existence of multiple words. If the universe computes the entirety of the wave function, including all of its multiple worlds, then all of the observers in those worlds should experience physical reality by the act of being computed—it is not at all clear how the classical particles could have physical reality and the rest of the wave function not.[21][42] In the words of David Deutsch, "pilot-wave theories are parallel-universes theories in a state of chronic denial. This is no coincidence. Pilot-wave theories assume that the quantum formalism describes reality. The multiplicity of reality is a direct consequence of any such theory."[21]

However, since the extra classical particles only exist in one of these worlds, the pilot wave interpretation also does not resolve the problem of the low likelihood of the observed state of the universe[10] or the complexity of the required initial condition[36]. Thus, the pilot wave interpretation, despite being strictly more complicated than the Everett interpretation—both in terms of its extra postulate and the concerns above—produces exactly no additional explanatory power. Therefore, we can safely dismiss the pilot wave interpretation on the grounds of the same simplicity argument used against the Copenhagen interpretation in section IV B.

VI. CONCLUSION

Harvard theoretical physicist Sidney Coleman uses the following parable from Wittgenstein as an analogy for the interpretation of quantum mechanics: "'Tell me,' Wittgenstein asked a friend, 'why do people always say, it was natural for man to assume that the sun went round the Earth rather than that the Earth was rotating?' His friend replied, 'Well, obviously because it just looks as though the Sun is going round the Earth.' Wittgenstein replied, 'Well, what would it have looked like if it had looked as though the Earth was rotating?"'[49] Of course, the answer is *it would have looked exactly as it actually does!* To our fallible human intuition, it seems as if we are seeing the sun rotating around the Earth, despite the fact that what we are actually seeing is a heliocentric solar system. Similarly, it seems as if

we are seeing the wave function randomly collapsing around us, despite the fact that this phenomenon is entirely explained just from the wave equation, which we already know empirically is a law of nature.

It is perhaps unfortunate that the Everett interpretation ended up implying the existence of multiple worlds, since this fact has led to many incorrectly viewing the Everett interpretation as a fanciful theory of alternative realities, rather than the best, simplest theory we have as of yet for explaining measurement in quantum mechanics. The Everett interpretation's greatest virtue is the fact that it is barely even an interpretation of quantum mechanics, holding as its most fundamental principle that the wave equation can interpret itself. In the words of David Wallace: "If I were to pick one theme as central to the tangled development of the Everett interpretation of quantum mechanics, it would probably be: the formalism is to be left alone. What distinguished Everett's original paper both from the Dirac-von Neumann collapse-of-the-wavefunction orthodoxy and from contemporary rivals such as the de Broglie-Bohm theory was its insistence that unitary quantum mechanics need not be supplemented in any way (whether by hidden variables, by new dynamical processes, or whatever)."[11]

There is a tendency of many physicists to describe the Everett interpretation simply as one possible answer to the measurement problem. It should hopefully be clear at this point why that view should be rejected—the Everett interpretation is not simply yet another solution to the measurement problem, but rather a straightforward conclusion of quantum mechanics itself that shows that *the measurement problem should never have been a problem in the first place.* Without the Everett interpretation, one is forced to needlessly introduce complex, symmetry-breaking, empirically-unjustifiable postulates—either wave function collapse or pilot wave theory—just to explain what was *already explicable* under basic wave mechanics. The Everett interpretation is not just another possible way of interpreting quantum mechanics, but a necessary component of any quantum theory that wishes to explain the phenomenon of measurement in a natural way. In the words of John Wheeler, Everett's thesis advisor, "No escape seems possible from [Everett's] relative state formulation if one wants to have a complete mathematical model for the quantum mechanics that is internal to an isolated system. Apart from Everett's concept of relative states, no self-consistent system of ideas [fully explains the universe]."[6]

[1] W. Heisenberg, The actual content of quantum theoretical kinematics and mechanics, Zeitschrift für Physik (1927).

[2] The solvay conference, probably the most intelligent picture ever taken, 1927.

[3] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete?, Physical Review (1935).

[4] D. M. Greenberger, Bell's theorem without inequalities, American Journal of Physics (1990).

[5] J. Townsend, *Quantum Physics: A Fundamental Approach to Modern Physics* (University Science Books, 2010).

[6] J. A. Wheeler, Assessment of everett's "relative state" formulation of quantum theory, Reviews of Modern Physics (1957).

[7] H. Everett, The theory of the universal wave function, Princeton University Press (1957).

[8] H. Everett, "relative state" formulation of quantum mechanics, Reviews of Modern Physics (1957).

[9] B. S. DeWitt, Quantum mechanics and reality, Physics Today (1970).

[10] A. Barrau, Testing the everett interpretation of quantum mechanics with cosmology (2015).

[11] D. Wallace, Quantum probability from subjective likelihood: Improving on deutsch's proof of the probability rule, Studies in History and Philosophy of Science (2007).

[12] S. Saunders, J. Barrett, A. Kent, and D. Wallace, *Many Worlds?: Everett, Quantum Theory, & Reality* (Oxford University Press, 2010).

[13] D. Wallace, *The Emergent Multiverse* (Oxford University Press, 2014).

[14] D. Wallace, Epistemology quantized: circumstances in which we should come to believe in the everett interpretation, The British Journal for the Philosophy of Science (2006).

[15] M. Arndt, O. Nairz, J. Vos-Andreae, C. Keller, G. van der Zouw, and A. Zeilinger, Wave-particle duality of c60 molecules, Nature (1999).

[16] M. C. Price, The everett faq (1995).

[17] S. S. Hawking, Black holes and thermodynamics, Physical Review D (1975).

[18] D. Deutsch, Quantum theory of probability and decisions, Proceedings of the Royal Society of London (1999).

[19] D. Wallace, Everettian rationality: defending deutsch's approach to probability in the everett interpretation, Studies in History and Philosophy of Science (2003).

[20] C. Clark, A theoretical introduction to wave mechanics (2010).

[21] D. Deutsch, Comment on lockwood, The British Journal for the Philosophy of Science (1996).

[22] S. Carroll, The wrong objections to the many-worlds interpretation of quantum mechanics (2015).

[23] J. B. Hartle, Spacetime quantum mechanics and the quantum mechanics of spacetime (2014).

[24] H. D. Zeh, Feynman's interpretation of quantum theory, The European Physical Journal (2011).

[25] S. Carroll, Why the many-worlds formulation of quantum mechanics is probably correct (2014).

[26] A. I. M. Rae, Everett and the born rule, Studies in History and Philosophy of Science (2009).

[27] R. J. Solomonoff, A preliminary report on a general theory of inductive inference (1960).

[28] A. N. Soklakov, Occam's razor as a formal basis for a physical theory (2001).

[29] A. Altair, An intuitive explanation of solomonoff induction (2012).

[30] L. Kelion, Coder creates smallest chess game for computers (2015).

[31] M. Bitbol, The concept of measurement and time symmetry in quantum mechanics, Philosophy of Science (1988).

[32] E. Yudkowsky, The quantum physics sequence: Collapse postulates (2008).

[33] J. Ellis and J. S. Hagelin, Search for violations of quantum mechanics, Nuclear Physics (1984).

[34] J. Ellis, J. L. Lopez, N. E. Mavromatos, and D. V. Nanopoulos, Precision tests of cpt symmetry and quantum mechanics in the neutral kaon system, Physical Review D (1996).

[35] M. Agrawal, Linearity in quantum mechanics (2003).

[36] H. D. Zeh, Measurement in bohm's versus everett's quantum theory, Foundations of Physics (1988).

[37] W. H. Zurek, Decoherence and the transition from quantum to classical—revisited, Los Alamos Science (2002).

[38] M. Schlosshauer, Decoherence, the measurement problem, and interpretations of quantum mechanics (2005).

[39] G. Bacciagaluppi, *The Role of Decoherence in Quantum Mechanics* (Stanford Encyclopedia of Philosophy, 2012).

[40] D. Wallace, Everett and structure, Studies in History and Philosophy of Science (2003).

[41] H. D. Zeh, On the interpretation of measurement in quantum theory, Foundations of Physics (1970).

[42] H. D. Zeh, Why bohm's quantum theory?,

Foundations of Physics Letters (1999).

[43] R. B. Griffiths, Consistent histories and the interpretation of quantum mechanics, Journal of Statistical Physics (1984).

[44] M. Gell-Mann and J. B. Hartle, Quantum mechanics in the light of quantum cosmology, Int. Symp. Foundations of Quantum Mechanics (1989).

[45] P. Wallden, Contrary inferences in consistent histories and a set selection criterion (2014).

[46] S. Lloyd and O. Dreyer, The universal path integral, Quantum Information Processing (2015).

[47] D. J. Bohm and B. J. Hiley, The de broglie pilot wave theory and the further development of new insights arising out of it, Foundations of Physics (1982).

[48] H. R. Brown and D. Wallace, Solving the measurement problem: De broglie-bohm loses out to everett, Foundations of Physics (2005).

[49] S. Coleman, Quantum mechanics in your face (1994).